Cover Sheet: Request 14331

STA4XXX Statistical Learning in R

Info

Process	Course|New|Ugrad/Pro
Status	Pending at PV - University Curriculum Committee (UCC)
Submitter	Michael Daniels daniels@ufl.edu
Created	10/10/2019 1:38:08 PM
Updated	11/17/2019 2:28:24 PM
Description of request	This is the second of three new statistics courses for the new data science major

Actions

Step	Status	Group	User	Comment	Updated
Department	Approved	CLAS - Statistics 011623000	Michael Daniels		10/11/2019
No document changes					
College	Conditionall Approved	CLAS - College of Liberal Arts and Sciences	Joseph Spillane	The College Curriculum Committee conditionally approves this request, with the following: add "in R" to the course title and transcript title; 2) add a 15th week to the weekly schedule of courses	11/16/2019
No document changes					
Department	Approved	CLAS - Statistics 011623000	Michael Daniels	made all requested changes except for adding 'in R' to transcript title due to character limits	11/16/2019
No document changes					
College	Approved	CLAS - College of Liberal Arts and Sciences	Joseph Spillane		11/17/2019
No document changes					
University Curriculum Committee	Pending	PV - University Curriculum Committee (UCC)			11/17/2019
No document changes					
Statewide Course Numbering System					
No document changes					
Office of the Registrar					
No document changes					
Student Academic Support System					
No document changes					
Catalog					
No document changes					
College Notified					
No document changes					

Course|New for request 14331

Info

Request: STA4XXX Statistical Learning in R
Description of request: This is the second of three new statistics courses for the new data science major
Submitter: Michael Daniels daniels@ufl.edu
Created: 12/6/2019 9:54:29 PM
Form version: 5

Responses

Recommended Prefix STA
Course Level 4
Course Number XXX
Category of Instruction Advanced
Lab Code None
Course Title Statistical Learning in R
Transcript Title Statistical Learning
Degree Type Baccalaureate

Delivery Method(s) On-Campus
Co-Listing No

Effective Term Fall
Effective Year 2020
Rotating Topic? No
Repeatable Credit? No

Amount of Credit 3

S/U Only? No
Contact Type Regularly Scheduled
Weekly Contact Hours 3
Course Description Overview of the field of statistical learning. Topics include linear regression, classification, resampling methods, shrinkage approaches, tree-based methods, support vector machines, and clustering. Approaches will be illustrated in R.

Prerequisites STA 4322 & STA 4210 & MAS 4XXX (Linear Algebra for Data Science)

Co-requisites None
Rationale and Placement in Curriculum This will be a core course for the new data science major (within statistics) and an elective within the statistics major. This course will introduce majors to the advanced methodological toolkit essential for understanding large and complex data that has arisen in fields from biology to finance to astronomy.

Course Objectives - Explain the theoretical underpinnings of the statistical learning methods

- Implement methods for classification in R

- Implement methods for regularization in R

- Implement statistical machine learning methods in R

Course Textbook(s) and/or Other Assigned Reading James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013) An Introduction to Statistical learning with Applications in R. Springer.

Weekly Schedule of Topics Week 1
- Introduction to Statistical Learning

Week 2
- Review of linear regression

Week 3
Classification
- logistic regression

Week 4
Classification (continued)
- linear discriminant analysis
- quadratic discriminant analysis

Week 5
Resampling methods
- Cross-validation

Week 6
Resampling methods (continued)
- k-Fold cross-validation
- Exam 1

Week 7
Linear model selection and regularization
- Subset selection

Week 8
Linear model selection and regularization (continued)
- Shrinkage methods
- ridge regression
- the lasso

Week 9
Linear model selection and regularization (continued)
- Dimension reduction methods
- principal components regression
- partial least squares

Week 10
Nonlinear models
- polynomial regression
- regression splines
- smoothing splines

Week 11
Tree-based methods
- Decision trees

Week 12
Tree-based methods (continued)
- Bagging
- Random forests
- Boosting
Exam 2

Week 13
Support Vector Machines
- Maximal margin classifier
- support vector classifier
- support vector machines

Week 14
Unsupervised learning
- Principal Components analysis

Week 15
Unsupervised learning (cont.)
- K-means clustering
- Hierarchical clustering

Grading Scheme Grades will be based on weekly homeworks and three in class exams.

Homeworks   10%
Exam 1   30%
Exam 2   30%
Exam 3   30%

Grades will be assigned as follows: 92.5-100, A; 90.0-92.4, A-; 87.5-89.9, B+; 82.5-87.4, B; 80.0-82.4, B-; 77.5-79.9, C+; 72.5-77.4, C; 70.0-72.4, C-; 67.5-69.9, D+; 62.5-67.4, D; 60.0-62.4, D-; 0-59.9, F
The numeric scores will be rounded to the nearest tenth.

Instructor(s) to be determined
Attendance & Make-up Yes
Accomodations Yes
UF Grading Policies for assigning Grade Points Yes
Course Evaluation Policy Yes